Comparison of fMRI signals related to perceived and memorized emotional faces

Tuuli Maaria Luostarinen

Master's Thesis

Psychology

Faculty of Medicine

July 2018

Supervisor: Viljami Salmela

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto Fakultet/Sektion – Faculty	Laitos/Institution– Department
Faculty of medicine/the department of psychology and logopedics	

Tekijä/Författare – Author
Maaria Luostarinen

Työn nimi / Arbetets titel – Title
Comparison of fMRI signals related to perceived and memorized emotional faces

Oppiaine /Läroämne – Subject
Psychology

Työn laji/Arbetets art – Level	Aika/Datum – Month and year	Sivumäärä/ Sidoantal – Number of pages
Master's thesis	07/2018	24

Tiivistelmä/Referat – Abstract

Visual working memory refers to the cognitive system responsible for the short-term storage and manipulation of visual information. Prevailing view suggests that there is a trade-off between memory capacity and precision: we can hold more items in memory with lower precision or fewer items with higher precision. Recent functional magnetic resonance imaging (fMRI) studies suggest distributed visual working memory representations in multiple brains areas from sensory visual to the parietal and frontal cortex.

This thesis consists of a visual working memory fMRI study, using human faces as stimuli. The purpose of this thesis was to examine the visual working memory precision for angry, neutral and happy faces and the memory representations in the face network, and to directly compare the neural activity while participants discriminated and memorized faces.

The participants discriminated and remembered faces precisely and were highly aware of occasional memory lapses, as shown by the confidence ratings of responses. In the primary visual cortex (V1) and fusiform face area (FFA), happy faces elicited higher brain activation than angry or neutral faces. The multivariate analysis of fMRI activation patterns showed correlations between the perception and memory tasks in these areas. Overall, the activations and correlations were higher in the right hemisphere, as expected. The correlations between perception and memory conditions were surprisingly low given the identical stimuli in these conditions. Even small positive correlations in the right V1 and FFA, however, support their role in maintaining facial information in visual working memory.

Avainsanat – Nyckelord – Keywords
Visual working memory, capacity, precision, fMRI, Face network

Säilytyspaikka – Förvaringställe – Where deposited
University of Helsinki – Helda / E-thesis (dissertations), *ethesis.helsinki.fi*

Muita tietoja – Övriga uppgifter – Additional information

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta/Osasto Fakultet/Sektion – Faculty	Laitos/Institution– Department
Lääketieteellinen tiedekunta/psykologian ja logopedian osasto	

Tekijä/Författare – Author
Maaria Luostarinen

Työn nimi / Arbetets titel – Title
Havaittujen ja muistettujen emotionaalisten kasvojen fMRI-signaalien vertaaminen

Oppiaine /Läroämne – Subject
Psykologia

Työn laji/Arbetets art – Level	Aika/Datum – Month and year	Sivumäärä/ Sidoantal – Number of pages
Pro gradu --tutkielma	07/2018	24

Tiivistelmä/Referat – Abstract
Visuaalisella työmuistilla tarkoitetaan kognitiivista järjestelmää, joka vastaa lyhytaikaisesta näönvaraisen tiedon säilyttämisestä ja muokkaamisesta. Vallitsevan näkemyksen mukaan kapasiteetin ja tarkkuuden välillä on aina tehtävä kompromissi: voimme muistaa enemmän asioita huonommalla tarkkuudella tai vähemmän asioita paremmalla tarkkuudella. Viimeaikaisten, toiminnallisella magneettikuvauksella (fMRI), tehtyjen tutkimusten mukaan visuaalisessa työmuistissa olevat ärsykkeet ovat jakautuneet usealle aivoalueelle, sensoriselta ja aivokuorelta lähtien päälaenlohkoon ja etuotsalohkoon.

Tämä tutkielma koostuu kasvoärsykkeillä tehdystä visuaalisen työmuistin fMRI-tutkimuksesta. Tarkoituksena oli tutkia visuaalisen työmuistin tarkkuutta iloisille, neutraaleille ja vihaisille kasvoille ja niiden muistirepresentaatioita kasvoverkostossa, sekä suoraan verrata erottelu- ja muistitehtävien aikaista aktivaatiota.

Itsevarmuusasteikon vastauksista päätellen, koehenkilöt kykenivät erottelemaan ja muistamaan kasvoja tarkasti ja olivat tietoisia muistinsa satunnaisista virheistä. Primaarilla visuaalisella aivokuorella (V1) ja fusiformisella kasvoalueella (FFA), iloiset kasvot aiheuttivat korkeamman aktivaation kuin vihaiset tai neutraalit kasvot. Monimuuttujaisissa analyyseissa fMRI-aktivaatiokuvioista löytyi muisti- ja havaintotehtävän välisiä korrelaatioita näiltä alueilta. Aktivaatiot ja korrelaatiot olivat ylipäätään korkeammat oikealla hemisfäärillä, odotusten mukaisesti. Havainto- ja muistitilanteiden korrelaatio oli yllättävän pieni, ottaen huomioon, että tilanteissa käytettiin identtisiä ärsykkeitä. Pienetkin positiiviset korrelaatiot V1:llä ja FFA:lla kuitenkin tukevat näiden alueiden osallistumista muistissa säilyttämiseen visuaalisessa työmuistissa.

Avainsanat – Nyckelord – Keywords
Visuaalinen työmuisti, kapasiteetti, tarkkuus, fMRI, kasvoverkosto

Säilytyspaikka – Förvaringställe – Where deposited
Helsingin yliopisto – Helda / E-thesis (opinnäytetyöt), *ethesis.helsinki.fi*

Muita tietoja – Övriga uppgifter – Additional information

Contents

1 Introduction 1
1.1 Visual working memory 1
1.2 Neuroimaging studies of visual working memory 3
1.3 Neural correlates of face perception 4
1.4 Working memory and faces 5
1.5 The present thesis 6
2 Materials and Methods 6
2.1 Participants 6
2.2 Stimuli 7
2.3 Procedure 8
2.4 fMRI acquisition and pre-processing 9
2.5 fMRI estimation 9
2.6 Region of Interest analysis 10
2.7 Multivariate analysis 10
3 Results 11
3.1 Behavioral data 11
3.2 Univariate fMRI 14
3.3 ROIs 14
4 Discussion 18
4.1 limitations 20
4.2 Conclusions 21
References 22

1 Introduction

1.1 Visual working memory

Working memory refers to the cognitive system responsible for the short-term storage and manipulation of information (Baddeley & Hitch, 1974; Baddeley, 1992). According to the classical model, it is divided into central executive, visual and auditory storages (Baddeley, 1992) as well as to episodic buffer (Baddeley, 2000). This thesis concentrates on the visual part of working memory or visual working memory, and especially to the contents in the visual memory storage and the neural locus of memory representations.

For several decades, the primary focus of studies on visual working memory has been its limited capacity (e.g., Cowan, 2001). The limited capacity means that only a certain number of items at a time can be stored in the working memory system (Luck & Vogel, 1997). It is usually defined as the discontinuity point in memory task performance (Wilken & Ma, 2004). The limit has been traditionally explained with a slot model, in which the memory storage consists of discrete and separate slots (Cowan, 2001). Each of these slots can contain only one stimulus at a time, so the capacity of the visual working memory is equal to the number of these slots (Cowan, 2001). Usually, the number of the slots has been thought to be 4±2.

However, 14 years ago, Wilken and Ma (2004) provided an alternative explanation for the discontinuity in the memory task performance. According to them, the discontinuity is due to a ceiling effect, meaning that the memory tasks with up to four stimuli are too easy, and the decrease in performance is only observed after the number of stimuli is raised so high that the ceiling effect is avoided. They also showed that when the memory task is made more difficult, the decline is steady from one stimulus onwards, and no clear discontinuity point can be found.

Wilken and Ma's (2004) study suggests that the precision of working memory might be more relevant than its capacity for understanding memory limitations. Memory precision is usually studied with a method of adjustment, recall or reproduction of the remembered stimuli on a continuum, for example, the 360º rotation of a line. In these studies, a target stimulus, which is a specific point in the continuum (The particular angle of the line), is first presented to the subject and their task is to memorize it. After a memory delay period, a probe stimulus from a random point of the continuum is shown, and the subject's task is to adjust the probe stimulus

to match the target stimulus. Precision is defined by how close to the target stimulus the subject adjusted the probe.

Precision of visual working memory has been studied with both simple visual features, for example orientation or colours (e.g., Salmela & Saarinen, 2013; Zhang & Luck, 2008; Bays & Husain, 2008) and complex objects for example faces (e.g., Jiang, Shim & Makovski, 2008; Lorenc et al., 2014). These studies concluded that there is a trade-off between the capacity and precision of visual working memory: we can hold more items in memory with lower precision or fewer items with higher precision.

Thus, an alternative to the slot model for working memory is a resource model, in which the overall capacity is also limited but instead of discrete slots, the capacity consists of a continuous resource that can be flexibly divided across memorized items (Bays & Husain, 2008). The resource model has been implemented as a neural population coding model, and it has been shown that neural noise in the population accounts for the errors in memory (Bays, 2014). This model predicts that the memory precision decreases steadily with the increase of the memory load, and there is no discontinuity point, which is in line with the previously mentioned results by Wilken and Ma (2004).

Zhang and Luck introduced the third popular visual working memory model in 2008. They separated the effects of precision and capacity using a mixture model. The basic idea of this model is that when the target stimulus is stored in memory, the subject will adjust the test stimulus close to the target stimulus and when it is not stored in memory, the subject will adjust the test stimulus randomly. When the distributions of both situations are added together, they create a mixture, which should represent the subjects' overall performance in the memory task. According to them, the visual working memory storage acts like an all-or-none process, the stimuli are either stored in the slots or not stored at all. Zhang and Luck suggested slots + averaging model, which can explain the changes in precision but still has discrete slots. In this model, one stimulus can be stored in multiple slots and thus increase precision by reporting the average of these representations. They concluded that the visual working memory capacity consists of three discrete slots, thus when the number of items exceeds them, only three of them are held in memory, and the averaging explains the differences in precision when the number of items is less than the number of slots.

Several studies have since made a comparison of these models. Jiang, Shim, and Makovski (2008) argued that their results from comparing faces and gratings, could be seen as support

for either of these models, but the precision of working memory is not different for more complex stimuli. In a systematic comparison of different memory models, the resource model outperformed slot models (van den Berg, Awh & Ma, 2014), but the debate is still ongoing, and other models have also been presented (van den Berg, Shin, George & Ma, 2012).

1.2 Neuroimaging studies of visual working memory

Functional magnetic resonance imaging (fMRI) is a spatially accurate method of studying brain activation and therefore it is often used to define which areas of the brain are responsible for specific functions. When an area of the brain activates, for example, due to cognitive tasks, the amount of oxygen-free hemoglobin is increased in that area. fMRI scanner can detect changes in the magnetic properties of the hemoglobin in blood, and thus fMRI measures blood-oxygen-level-dependent (BOLD) signal. Because the BOLD signal is slow, in the fMRI analysis, it is modeled with the hemodynamic response function which is created by modeling the amount of hemoglobin as concentration over time. The BOLD signal is a secondary signal, and thus fMRI does not measure the neural activity itself but instead the blood flow in the brain.

When fMRI became popular at the beginning of this century, univariate methods were used to analyze the data (Lee & Baker, 2016). When using the univariate method, the data is approached by modeling the BOLD data with a general linear model (GLM) within each voxel individually. The beta values of the regressors in the GLM are then tested, usually with a one-way t-test, by comparing them to a baseline condition (rest) or beta values during other stimuli or task. With univariate methods, sustained activity during the time that the stimuli are held in memory (memory delay), is observed in the frontal areas of the brain (e.g., Haxby, Petit, Ungleider, & Coyrtney, 2000).

Multivariate pattern analysis (MVPA) is a more recent method of analyzing fMRI data. The MVPA -method is more sensitive to detecting neural activation than the univariate methods (Kriegeskorte, 2011). In MVPA the pattern information or distribution of BOLD activity within a group of voxels is analyzed (e.g., Kriegeskorte, 2011; Haxby, Connolly, & Guntupalli, 2014). With the MVPA, stimuli representations during the memory delay have been decoded for example from the visual areas of the brain (for review see: Lee & Baker, 2016). It is still being debated, however, whether the storage part of the visual working memory is somewhere in the frontal-parietal -network (D'Esposito & Postle, 2014) or visual areas (Lee & Baker, 2016) of the brain.

In their review, Lee and Baker (2016) noted that because a visual working memory task requires attention, this could explain why the frontal areas of the brain also activate. Also, the ability to hold information could be a property of the whole cortex and not isolated to only specific areas. Christophel, Lamshchinina, Yan, Allefeld, and Haynes (2018) found that the attended stimuli are indeed held in the visual areas, the intraparietal areas, and the frontal eye fields, but the unattended stimuli are held only in the intraparietal regions and the frontal eye fields. They suggested that the visual areas were essential to achieving better memory precision for the attended stimulus.

In sum, visual working memory task is a highly straining task for the brain, thus requiring the co-operation of multiple different areas. That being said, according to current research, it is likely that the stimulus is held in the visual areas and the frontal regions are responsible for other cognitive functions needed during the memory tasks.

1.3 Neural correlates of face perception

The human face is an excellent example of a real-life, complex, visual stimulus. It contains information about both identity and expression and is, therefore, an essential part of socializing (Haxby, Hoffman & Gobbini, 2000). Face recognition has also been proposed as the most developed visual perception skill in humans (Haxby, Hoffman & Gobbini, 2000). Humans have a build in preference for faces, and most humans spend more time in their life looking at faces than other objects (Haxby, Hoffman & Gobbini, 2000). Information on faces is processed in a brain network, containing multiple areas from the cortical regions of frontal and temporal lobes, other cortical areas and subcortical structures (Atkinson & Adolphs, 2011). From all these areas Atkinson and Adolphs (2011) highlighted the occipital face area (OFA) and fusiform face area (FFA) as the primary areas of face processing.

Facial identity and facial expression are thought to be recognized by different neural systems (Haxby, Hoffman, & Gobbini, 2000). Haxby and colleagues (2000) proposed a model in which the brain areas are divided to process either invariant (e.g., identity) or changeable (e.g., expression) aspects of the face. In their review, Calder and Young (2005) noted that many psychological studies had supported this division. For example, a healthy person can recognize an expression from both familiar and unfamiliar faces (Calder & Young, 2005). Also, there have been cases of brain injuries affecting either identity or expression recognition separately (Calder & Young, 2005). However, according to them, these neural systems are not entirely separate, but rather overlapping. Rhodes and colleagues (2015) found supporting

evidence to this claim of overlapping systems and added that there is individual variation in the ability to process identity and expression.

The core of the face network has been proposed to be the OFA, FFA and the superior temporal sulcus (STS) (Haxby, Hoffman, & Gobbini, 2000; Zhen, Liu, Song & Liu, 2017) with sometimes the addition of amygdala for the expression processing (Atkinson and Adolphs, 2011). The other areas related to face processing are part of the extended face network. Zhen and colleagues (2017) found in their study of people with prosopagnosia, a neurological disorder that impairs face recognition, that impaired functional connections between the core face network areas predict a behavioral deficit in face recognition.

The STS is especially important relating to facial movements (Schobert, Corradi-Dell'Acqua, Frühholz, Zwaag, & Vuilleumier, 2018). Schobert and colleagues (2018) found in their fMRI study that distinct parts of the STS were responsible for distinct movements of the face (speech and expression). In univariate analysis, they found that the activations associated with different movements overlapped largely, but with MVPA they were able to decode all conditions from the STS. They showed that univariate analyses are not sensitive enough to differentiate between neutral, angry and happy expressions, but that these can be found with MVPA from the STS.

1.4 Working memory and faces

Due to the properties mentioned above, the face is an interesting stimulus to study in the context of working memory. As indicated, faces are processed uniquely, which could mean that remembering them is also unique. However, Jiang and colleagues (2008) noticed in their research that the capacity of working memory was no different to simple objects (line orientation), compared to faces.

Towler, Kelly, and Eimer (2015) found in their study, in which they measured event-related potentials (ERP) that the capacity of visual working memory for faces is directly dependent on the attention process. They too described this phenomenon as a trade-off between the capacity and precision of visual working memory. They also hypothesized that the limited capacity of visual working memory could be dependent on the position dependence of the representations in the primary visual cortex since it is spatially organized. They also noted that in addition to being complex the faces share visual features, and this makes maintaining multiple representations in memory especially challenging.

Different areas, responsible for the neural processing of remembered faces have been suggested. LoPresti, Schon, Tricarico, Swisher, Celone, and Stern (2008) proposed that when it comes to working memory tasks with both identity and expression, the network responsible of processing this information consists of the orbitofrontal cortex, amygdala, and hippocampus. Jackson, Wolf, Johnston, Raymond, and Linden (2008) in turn found that angry and happy faces activated the core face network during a visual short-term memory task.

In conclusion, it has been proposed that the visual working memory precision is as accurate for face stimuli as to simple objects. The neural correlates for visual working memory task with faces are not concise across different studies. However, research with other visual stimuli suggests that the stimuli are held in storage during the memory delay, in the same areas that activate during visual perception, so it could be hypothesized that face stimuli are held in the face network.

1.5 The present thesis

Since there is evidence that the memory representations are held in the visual areas in visual working memory tasks, and there is a network of areas in the brain especially for face stimuli, it should be possible to find face related information during memory delay from these areas. Therefore, the present thesis focuses on examining the differences between perceived faces to memory representation in face network.

The purpose of this thesis is to examine the working memory precision for faces and to compare the memory representations in the brain during memory delay to the representations activated in a perception task. The expectation is that these memory representations are found within the face network and that the neural activity for them is similar, and thus correlated, to the neural activity of perceived faces.

2 Materials and Methods

2.1 Participants

In total 17 volunteers (age mean: 28, std: 4.47; nine males and eight females) participated in the study. Three of the participants were left-handed. The participants were recruited from

University of Helsinki's student organizations' e-mail lists. All participants had normal or corrected to normal vision. Participants received a monetary compensation (10€/h) for participating in the experiment. The experiment was approved by the University of Helsinki Ethical Review Board in the Humanities and Social and Behavioral Sciences.

2.2 Stimuli

The stimuli consisted of greyscale images of Caucasian male and female faces from the Faces database (Ebner, Riediger, & Lindenberg, 2010). From each identity, three different pictures were used: one neutral, one angry and one happy. These pictures were morphed to each other to form continuums from happy to neutral to angry. From this continuum 9 points were chosen to represent different intensities of the emotions: Neutral, 8 % angry/happy, 16 % angry/happy, 24 % angry/happy, and 40 % angry/happy. Figure 1 shows examples of these stimuli.

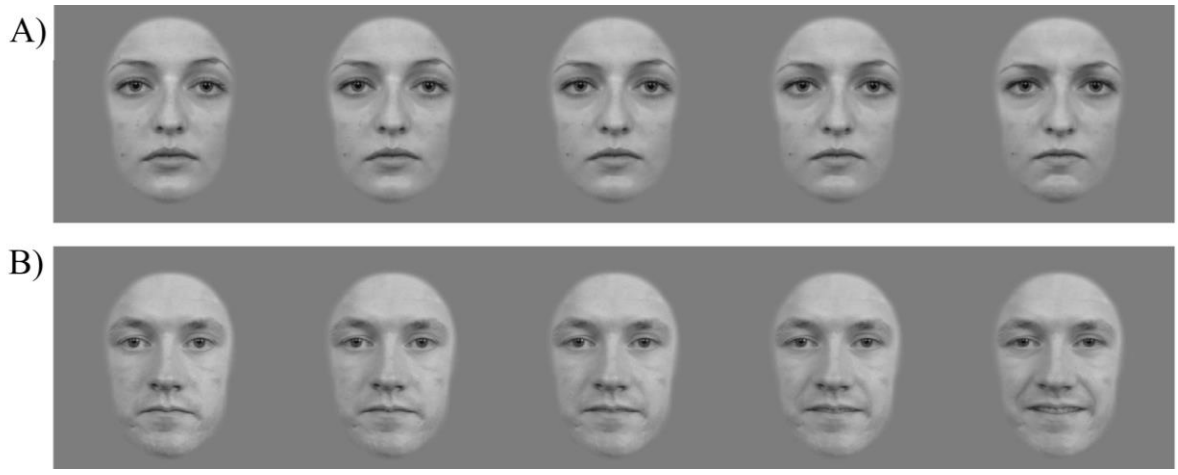

Figure 1. *Examples of the stimuli. A) Angry intensities, from left to right: neutral, 8 %, 16 %, 24 % and 40 % B) Happy intensities, from left to right: neutral, 8 %, 16 %, 24 % and 40 %.*

In addition, checkerboards, phase-scrambled faces, and pictures of objects were used to localize visual areas. Display size was 1920 × 1200 pixels (45.8º × 29.6º), the faces were 416 × 615 pixels (10.4º × 15.4º). The faces were presented at the center of the screen on a grey background. The participant viewed the stimuli at 40 cm distance.

2.3 Procedure

In the first experimental condition, a psychometric function for discriminating facial expressions was measured with a method of constant stimuli. On each trial, an image of a face was presented at the center of the screen on a grey background. The stimulus was present for 700 ms at a time. After the stimulus, a question appeared on display, and the participants' task was to answer if the expression on the face was angry or happy, on a 4-point confidence rating scale (certainly angry, maybe angry, maybe happy, certainly happy).

In the second experimental condition, a two-interval memory task was used. On each trial, an image of a face was presented to the participant, and they had to memorize it. After a ten-second memory delay, a second picture appeared on display, and the participant had to determine whether the emotion on the face had changed to angrier or happier, on a 4-point confidence rating scale (certainly angrier, maybe angrier, maybe happier, certainly happier). In half of the trials, the second face was happier and in half angrier than the first face. In 60 % of the trials, the change was ± 8 % and in 40 % ± 16 %.

Both conditions were done in separate runs. Each of the nine intensity levels was repeated four times in one perceptual run and two times in one memory run. Thus, each perceptual run consisted of 36 trials and memory runs 18 trials, and each intensity level was presented 24 times in total. Different expressions and intensity levels were presented in random order. There was a 30-second-long rest period in the middle of each run. Perception condition was repeated two times and the memory condition four times. There was a brief break between the runs. The participants practiced the tasks outside the scanner by completing one run of both conditions once.

In addition, a functional localizer run was measured. In each trial, one type of stimulus (face, checkerboard, horizontal checkerboard, vertical checkerboard, phase-scrambled face or picture of an object) was presented at the center of the screen. The participants were instructed to keep their gaze at the center of the display by looking at a cross. There were no other tasks in this condition. This condition consisted of 216 trials. Each type of stimulus was presented 36 times in one run. In total, the experiments lasted for 75 minutes.

2.4 fMRI acquisition and pre-processing

The fMRI scanner that was used was a 3 Tesla Siemens MAGNETOM Skyra at the AMI center of the Aalto University (Espoo, Finland), with a 30-channel head coil. The MR parameters were: echo time (TE)/repetition time (TR) = 30/2400 ms, Flip angle = 75º and the field of view (FOV) = 192 mm × 256 mm. 39 axial slices were acquired in the siemens-specific order (first the odd slices and then the even slices). The first two functional runs (perception) included 106 volumes, the next four runs (memory) included 124 volumes and the last one (localizer) included 182 volumes. All runs started with four empty scans (dummy scans) to ensure steady magnetization.

Pre-processing for the fMRI data was done by using the SPM12 (http://www.fil.ion.ucl.ac.uk/spm/) software package for MATLAB 8.0 (MathWorks Inc., Natick, Massachusetts, USA). The difference in slice timing of the BOLD images was corrected, and images realigned to the first image within a series. The functional EPI-images were co-registered and matched spatially to the anatomical images and corrected for movements. The images were then re-sliced and spatially smoothed by using a 6mm Gaussian filter. For the univariate analysis, the images were normalized by co-registering them into Montreal Neurological Institute (MNI) stereotactic space and resampled to isotropic voxels of 3 mm × 3 mm × 3 mm.

2.5 fMRI estimation

Univariate analysis consisted of series of contrast calculations. The regressors were created by modeling the BOLD data with a general linear model (GLM). For the first two functional runs, individual regressors were created to represent each nine intensity levels, and also, eight nuisance regressors were created (six movement parameters, question, and response). All the regressors, except for the movement parameters, were event-related and based on the stimuli/response onset. Regressors were convolved with the hemodynamic response function (HRF).

Finite impulse response (FIR) analysis was used to analyze the memory runs. The GLM was similar to the perception runs, except that each stimulus was modeled with five regressors, one for each time point of the memory maintenance and seven nuisance regressors were created (response and six movement parameters). In the FIR analysis, the HRF was not applied.

For the functional localizer run, regressors were created for all the different stimuli. Additionally, six nuisance regressors were created for this condition (movement parameters). All the regressors, except for the movement parameters, were event-related and based on the stimuli onsets.

After estimating the beta coefficient for each regressor, the differences between the mean beta values were tested with one-sample t-tests (univariate) or the distribution of beta values were compared (correlation, MVPA) in the second level analysis.

2.6 Region of Interest analysis

The regions of interest (ROIs) were chosen from the face network. The selected areas were: FFA, the primary visual cortex (V1), amygdala and the STS. ROI coordinates for FFA on both hemispheres were drawn from the regressors of the functional localizer by taking all the significant voxels from the second level one-sample t-test analyses. For FFA and the STS, the regressors used were face contrasted to phase-scrambled face. For FFA the minimum cluster size (k) was set to 10 and voxel-wise threshold at $p<.001$. Three significant clusters (cluster level $p<.001$), two of which were significant after Family Wise Error (FWE) correction, were found in the temporal lobe. For the STS the ROIs were drawn from the spmT images of these regressors. For both the V1 and amygdala, Anatomy toolbox for spm12 (Eickhoff, Paus, Caspers, Grosbras, Evans, Zilles, Amunts, 2007) was used to create ROIs representing these areas in both hemispheres. Anatomical V1 ROI was also overlapped with checkerboard localizer in order to select only voxels corresponding to the retinotopic location of the stimuli.

2.7 Multivariate analysis

For MVPA, representational similarity analysis (RSA, for review see: Kriegeskorte, Mur & Bandettini, 2008), was used. Representational dissimilarity matrices (RDMs) were computed from the activity patterns in the V1 and FFA during the perception and memory trials, so that one cell in the RDM represented the correlation of the response pattern for one intensity to the response pattern of one other intensity, thus all the RDMs were 9 x 9 matrices. Since some pattern information has proven to be lost in the process of smoothing the BOLD data (Kriegeskorte, Mur & Bandettini, 2008), unsmoothed data was used, and the ROIs were extracted individually from each participants' brain. Coordinates for the FFA on both hemispheres were drawn from the regressors of the functional localizer. The regressors used were face contrasted to phase-scrambled face, by taking all the significant voxels up to 100

voxels from the second level one-sample t-test analyses. For the V1 the coordinates were drawn from the functional localizer regressor, and the voxel limit was 500. These RDMs were then compared to each other in the RSA. The resulted correlations were then tested with a series of t-tests.

3 Results

3.1 Behavioral data

The perception task runs were analyzed by creating two psychometric functions across all participants, one for confident responses (certainly angry/happy) and one for uncertain responses (maybe angry/happy). This was done by calculating the frequency of responses from each intensity across all the participants and then fitting a cumulative normal distribution to those values. Figure 2 depicts the average psychometric function across all participants for confident and uncertain responses. Overall the participants recognized most of the emotions correctly. The neutral faces were more often rated as angry than happy. For uncertain trials, the discriminability was lower than for confident trials.

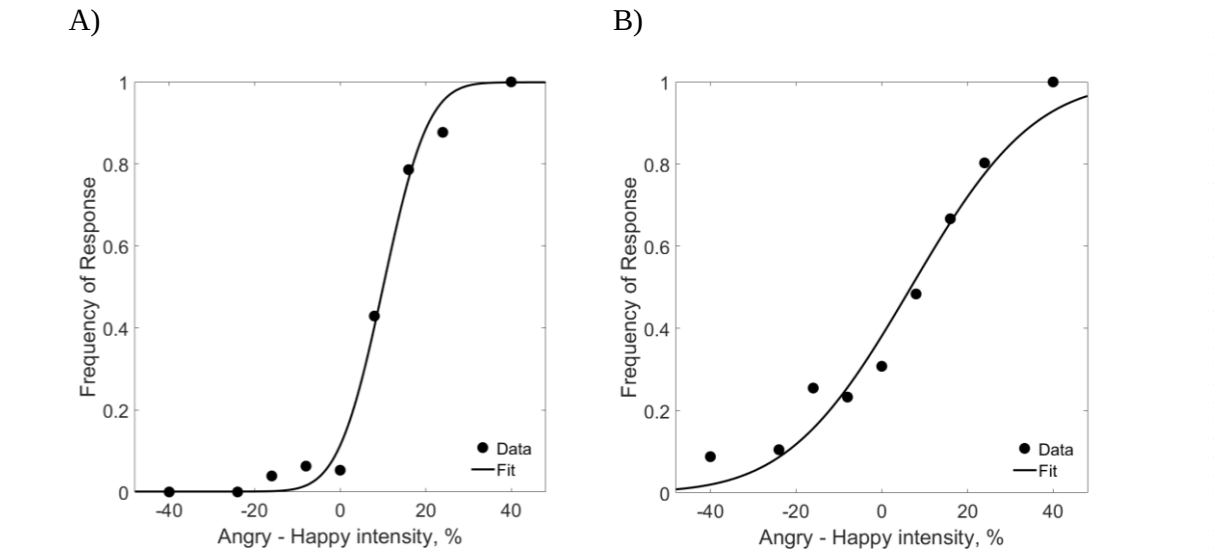

Figure 2. *A) The average psychometric function for confident responses across all participants in the perception task (mean=10.02, std=8.37, R^2=.99) B) The average*

psychometric function for uncertain responses across all participants in the perception task (mean=6.8, std=22.6, R^2=.97).

The number of correct responses in the memory task was 91,99 %. The participants responded with all confidence options equally. Figure 3 shows that when the change between the facial expressions was towards angrier, the participants responded with options one and two (certainly angrier and maybe angrier) more than options three and four (maybe happier and certainly happier) and vice versa.

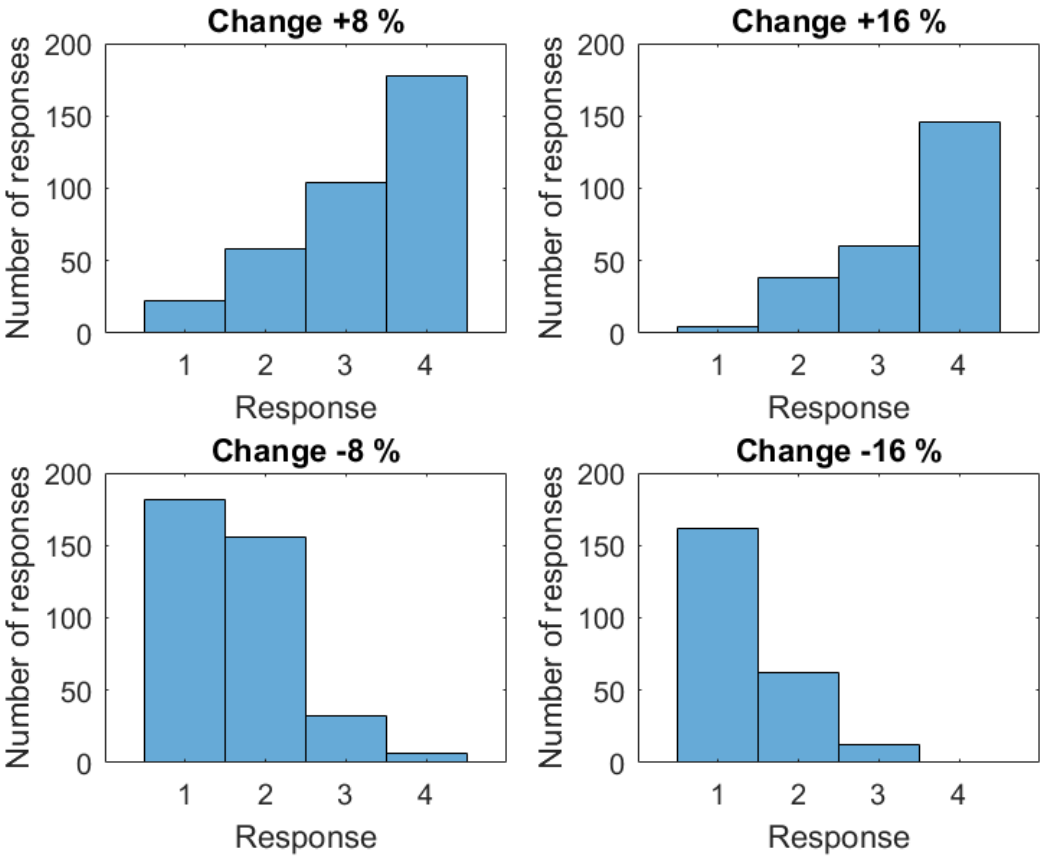

Figure 3. *The number of responses for each response option (1=certainly angrier, 2=maybe angrier, 3=maybe happier, 4=certainly happier), for each amount of change (+8 %: change eight percent towards happier, +16 %: change 16 percent towards happy, -8 %: change eight percent towards angry, -16 %: change 16 percent towards angry).*

The memory runs were also analyzed by creating two psychometric functions across all participants, one for confident responses (certainly angry/happy) and one for uncertain responses (maybe angry/happy). This was done by calculating the frequency of correct responses for angry and happy test faces across all the participants and then fitting a cumulative normal distribution to those values. The cases where the test face was neutral, were excluded from this analysis. Figure 4 shows the average psychometric functions for confident and uncertain responses in the memory trial. The standard deviation of certain responses being smaller (6.10) than that of uncertain responses (84.59) means that the participants were more often correct when using the confident responses than when using the uncertain ones. This suggests that participants had metacognitive knowledge on their memory representations.

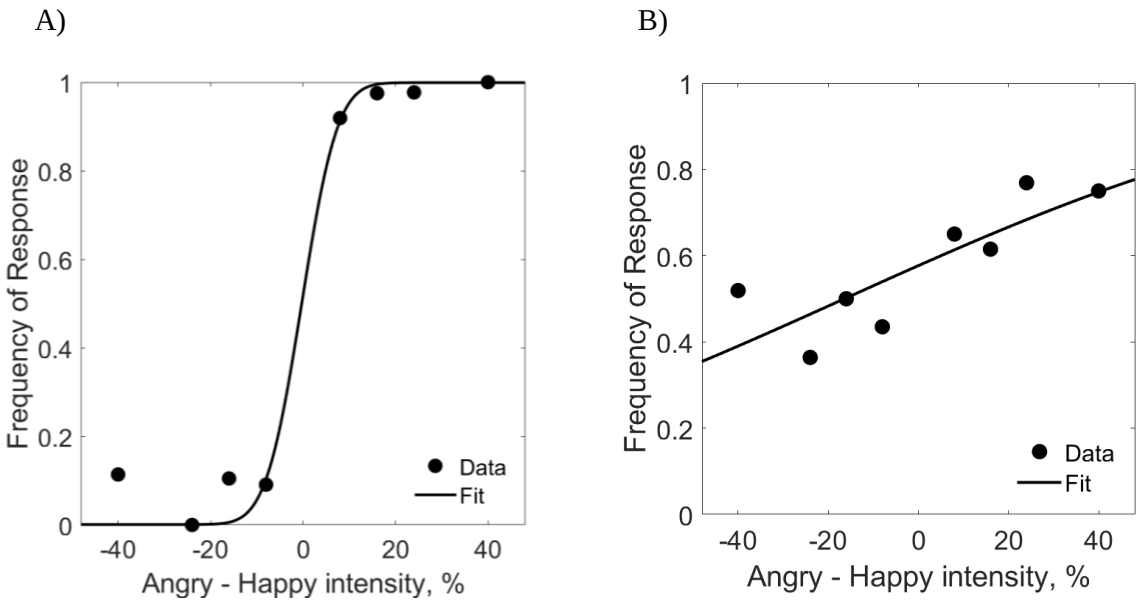

Figure 4. *A) The average psychometric function of confident responses across all participants (mean=-0.29, std=6.10, R^2=.99). B) The average psychometric function of uncertain responses across all participants (mean=-16.30, std=84.59, R^2=.68).*

3.2 Univariate fMRI

The face stimuli used in the experiment activated the face processing network (Figure 5, all regressors in perception condition summed). However, there were no significant differences between neutral, angry and happy faces in the univariate analyses, at the whole brain level.

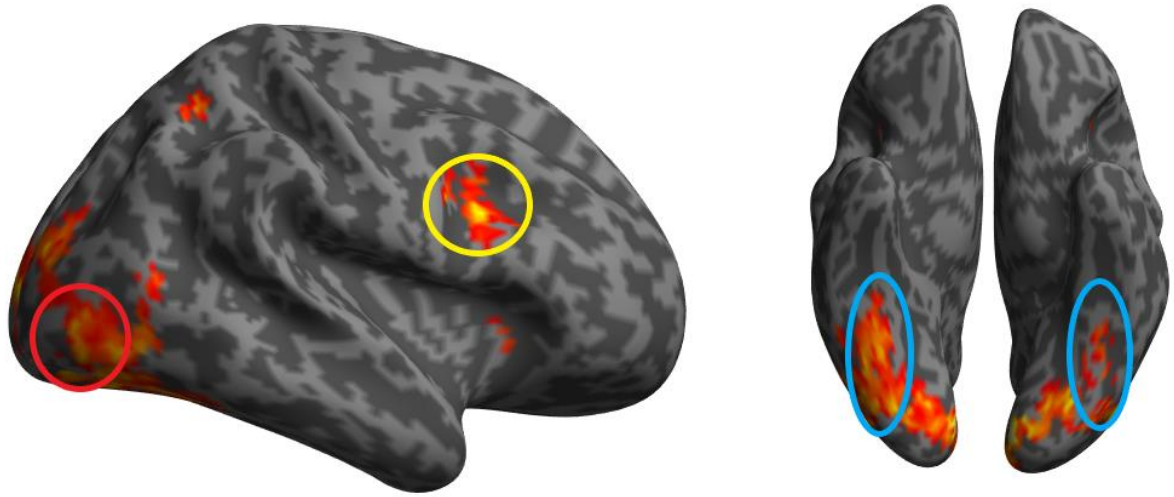

Figure 5. *Angry, happy and neutral faces activated the face network: the OFA (red), the middle frontal gyrus (yellow) and FFA on both hemispheres (blue).*

3.3 ROIs

Figure 6 shows the signal changes in the FFA, the V1, Amygdala, and the STS during the memory delay. As seen in the figure, the activity shows the typical pattern of a hemodynamic response in the FFA, the V1, and the STS. There was no sustained activation on any of the ROIs; the activity drops towards zero towards the end of the memory delay. On amygdala, the response is slower than on other ROIs, and the peak activation point is reached later during the memory delay.

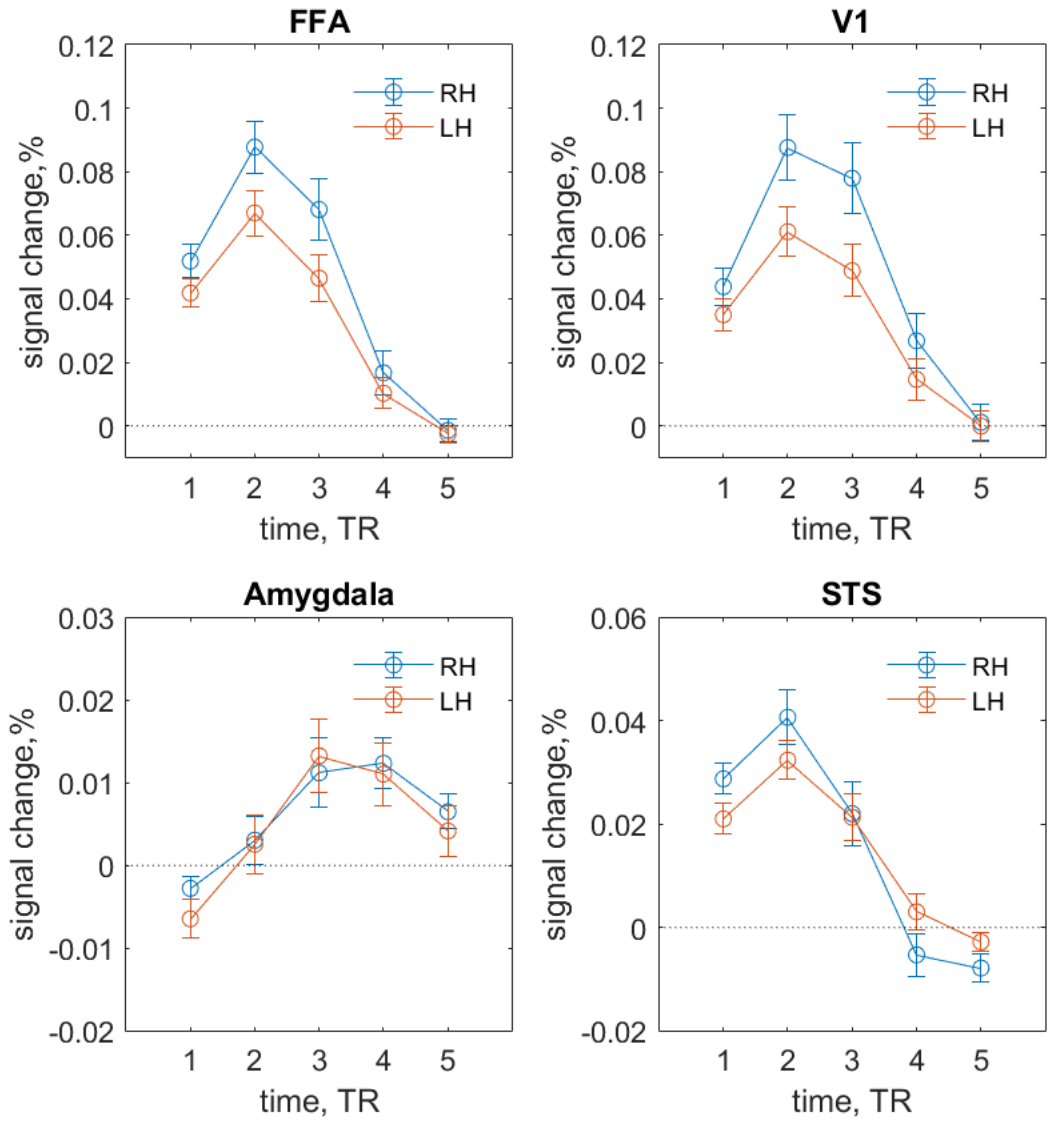

Figure 6. *Signal changes in FFA, V1, Amygdala, and STS during the memory delay between participants in the left hemisphere (LH) and the right hemisphere (RH) (the figures are scaled according to the individual maximum and minimum signal changes in the areas). The error bars represent the standard error means.*

Repeated measures ANOVA was done for the signal changes in all ROIs. Since the sphericity assumption was not fulfilled, Greenhouse-Geisser correction was used, but the degrees of freedom are reported according to the original ones, and the ones marked with ε are the corrected ones. Since the overall activity typically differs between brain regions, differences

between ROIs were not inspected further. Instead, the signal changes for different stimulus types within each ROI were analyzed.

Figure 7 shows the signal changes in all ROIs for each emotion. Only on STS, emotion had statistically significant effect ($F(2,32; \varepsilon 1.71,27.28)=3.96$, $p=.037$). On the STS, happy faces appear to have elicited overall larger responses. The Hemisphere × emotion interaction was statistically significant on FFA ($F(2,32; \varepsilon 1.68,26.91)=18.31$, $p<.001$), the V1 ($F(2,32; \varepsilon 1.75,28.07)=7.25$, $p=.004$) and STS ($F(2,32; \varepsilon 1.65,26.40)=10.15$, $p=0.001$), in all of these ROIs the emotions elicited larger responses in the right hemisphere.

Time point × emotion interaction was statistically significant on FFA ($F(8,128; \varepsilon 3.41,54.57)=8.99$, $p<.001$), the V1 ($F(8,128; \varepsilon 2.85,45.57)=7.03$, $p=.001$), amygdala ($F(8,112; \varepsilon 4.87,68.22)=3.38$, $p=.009$) and STS ($F(8,128; \varepsilon 3.9,62.39)=9.45$, $p<.001$). It seems that in FFA happy faces elicited stronger responses than neutral or angry faces during the end of the memory delay but not in the beginning. In the V1 the difference between conditions was most pronounced in intermediate time points. The conditions differed only in early (right) or late (left) phase of the memory period in the amygdala. In the STS it seems that happy faces differed from other stimuli mostly in the intermediate time points.

Hemisphere × time point × emotion interaction was statistically significant on FFA ($F(8,128; \varepsilon 4.67,74.74)=8.71$, $p<.001$), the V1 ($F(8,128; \varepsilon 5.59,89.49)=5.59$, $p<0.001$) and STS ($F(8,128; \varepsilon 5.46,87.37)=7.71$, $p<.001$). In FFA, right hemisphere had larger responses for all emotions at the beginning of the memory delay, but in addition, happy faces elicited a larger response in comparison to neutral and happy faces. In the V1 right hemisphere had overall larger responses at the beginning of the memory delay but also the emotions had larger signal differences. On the STS the responses on the right hemisphere were stronger on all emotions at the beginning of the memory delay.

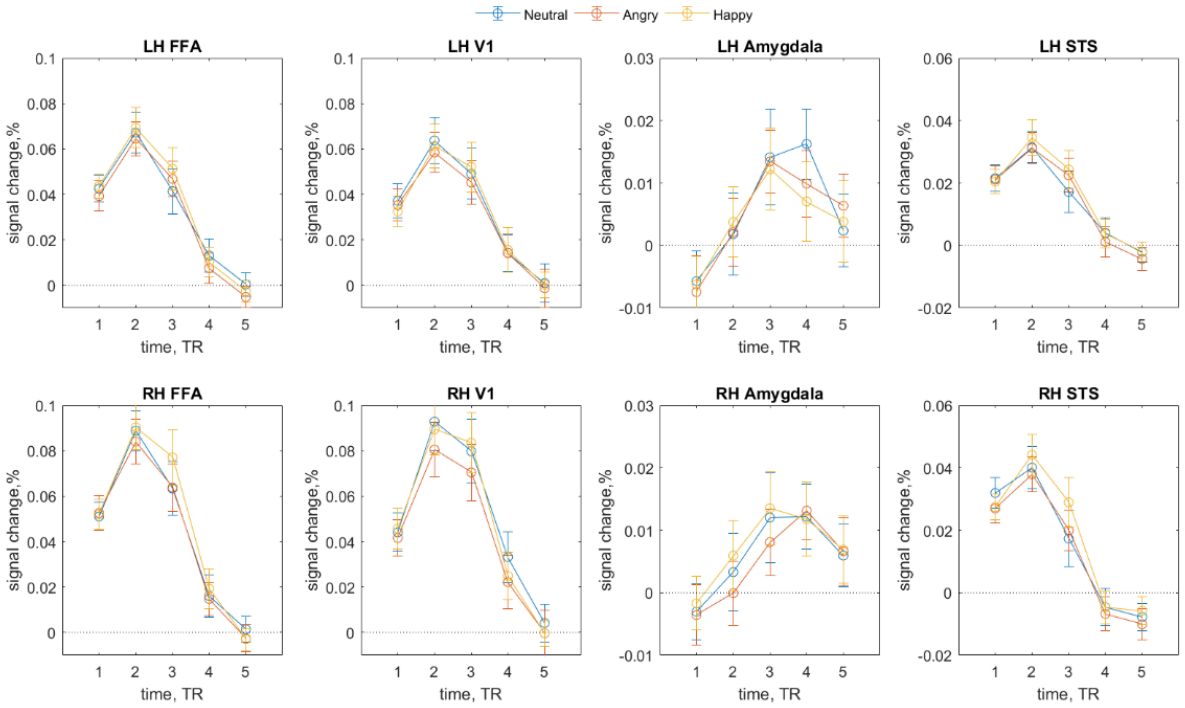

Figure 7. *Signal changes for each emotion in all ROIs from the left hemisphere (LH) and the right hemisphere (RH). The error bars represent the standard error means.*

3.4 Multivariate fMRI

Figure 8 displays the correlations of the RDMs for the perception runs to the RDMs for the memory runs within and between the V1 and FFA. The pattern information between the perception task and the memory task correlates positively in FFA during the total memory delay, suggesting that there is some sustained, face-related, activation in FFA. However, on the V1 the correlation rises with a delay but then drops. The same happens in the correlation of V1 perception task patterns to FFA memory task patterns, especially in the right hemisphere.

In a repeated measures ANOVA, neither hemisphere or time point had a statistically significant effect. The correlation of time point 3 in the V1 to V1 comparison in the right hemisphere differed from zero statistically significantly ($t(16)=2.96$, $p=0.009$) and also in the V1 to FFA comparison in the right hemisphere ($t(16)=3.12$, $p=0.007$). The differences between correlations in each hemisphere or the correlations of different time points compared to each other were not statistically significant.

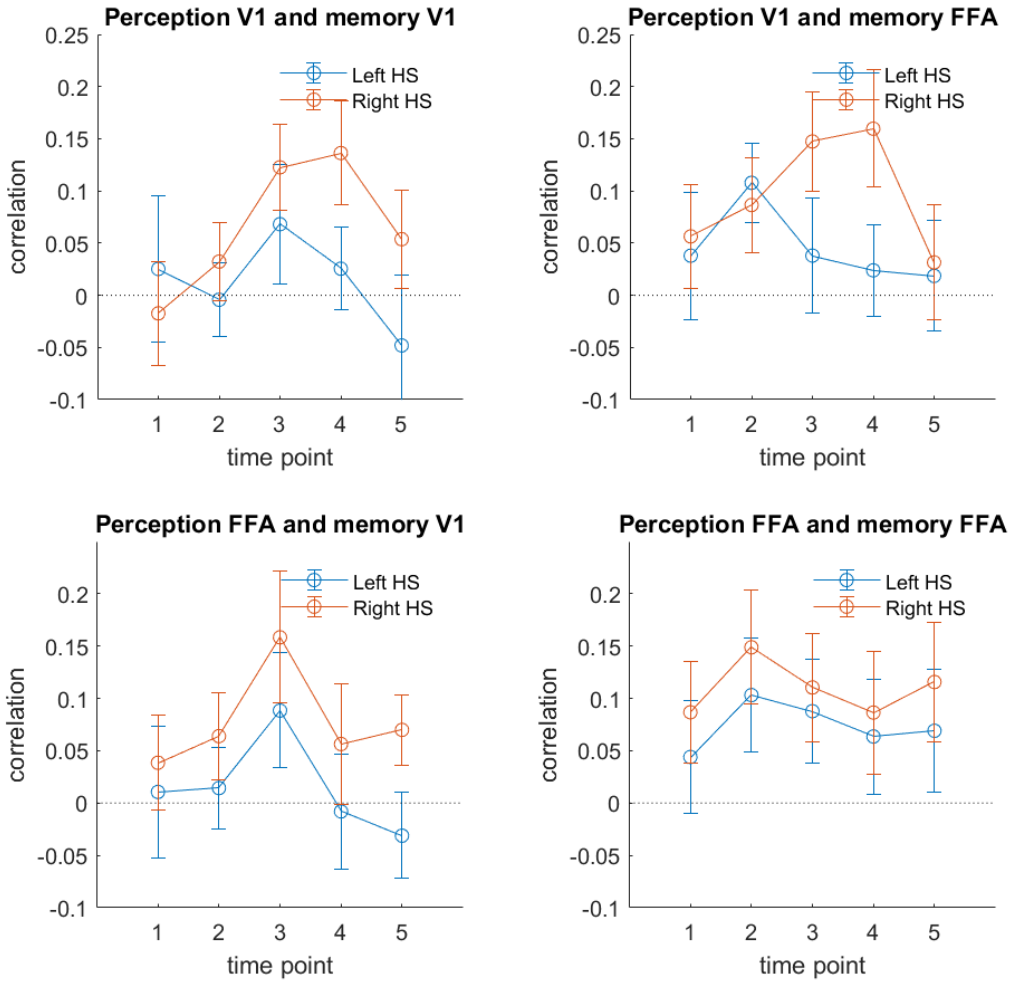

Figure 8. *Correlations of the RDMs from perception runs to the RDMs from memory runs in both hemispheres (HS) in the V1 and the FFA, the error bars represent the standard error means.*

4 Discussion

The purpose of this thesis was to examine the working memory precision for faces, to examine the memory representations in the brain during memory delay and compare these representations to the representations during face perception. The expectation was that brain areas involved in memory maintenance should show high correlations with perceptual representations. While robust activity was found in brain areas involved in face network during the memory task, the correlations between perceptual and memory representations

were relatively low and not statistically significant. However, all selected ROIs had some activation during the memory delay, and the signals were larger on the right hemisphere.

It appears that there is some sustained pattern activation in FFA during the memory delay, which suggests that there is a generic face representation in storage. The way that the correlation rises with a delay in V1 to V1 comparison could mean that the pattern information is more related to analyzing the facial features than holding a representation in memory.

The memory task itself seems to have been relatively easy since 91.99 % of the answers were correct. When evaluating the confidence ratings for the answers in the easy (change ±16 %) and the difficult (change ±8 %) conditions, the participants responded more confidently (using the options 1/4 more often than options 2/3 in their responses) in the easy memory condition, which seems rational since the larger the change is the easier it is to detect. When looking at the psychometric functions derived from the memory responses, it appears that when the participants were confident with their response (answered with the options 1=certainly angrier or 4=certainly happier), they were more likely to be correct than when answering with the uncertain options (2=maybe angrier or 3=maybe happier). It could be that the participants answered with more confidence when they were sure that they had remembered the test face with high precision, meaning that they had some metacognitive knowledge of their performance and the occasional lapses of their memory.

In line with the previous research, univariate analyses were not sensitive enough to separate between different emotions at the whole brain level (e.g., Schobert et al., 2018). However, in the ROI-analyses emotion and time point interaction had significant effects in all selected ROIs. Meaning that during the memory delay there were differences in the activation between angry and happy faces in the face network, mainly so that the happy faces elicited larger responses. This result differs from previous studies, where angry faces have shown higher activity and the memory task has even been easier with angry faces than happy faces (for example: Jackson, et al., 2008). This result could be affected by the fact that the used emotions did not represent the full emotions, since the highest intensities were 40 % happy or angry. It could also be that the participants perceived the happy faces as more pleasant and thus those representations were more actively maintained during the memory delay. Higher signal changes were overall observed in the right hemisphere. This is in align with previous studies showing that face processing is lateralized to the right hemisphere (Haxby, Hoffman & Gobbini, 2000).

The activity in most areas of the face network resembled the pattern of hemodynamic response and was not sustained. That does not exclude the possibility that the memory representations are held in these areas. Holding a face in memory does not necessarily require that constant high activity is maintained through the memory delay. Instead, the memory representations could be stored as activity patterns across the face network. Thus, instead of just examining the signal changes in the relevant areas, it is essential to study the spread of activity or activity patterns.

Since pattern information is often lost in the process of smoothing the BOLD data (Kriegeskorte, Mur & Bandettini, 2008), the MVPA was done for the unsmoothed data and the ROIs were extracted from each participants' own brain. From the results of these analyses it seems that in the right hemisphere areas V1 and FFA, some information on the face was preserved during the memory delay. Interestingly this information was more similar to the information during perceptual trials in V1 than in FFA. The pattern information during the perceptual task in FFA produced sustained correlation in FFA, and the correlations were above zero during the whole memory delay. Although robust activation in the face network was found, the correlation values between trials containing identical stimuli were modest. The participant count for this study was relatively small (n=17), so it could be that with a bigger sample size the correlations would have been higher.

4.1 Limitations

The fMRI in its' nature is a spatially accurate method and what is gained in this spatial accuracy is also lost in the accuracy of timing. In the future, it could be relevant to add other, more temporally accurate, measuring techniques (for example EEG) to the equation. Another limitation of fMRI is that the duration of the feasible experiment is practically limited to an hour. Because of this, the memory task couldn't be done with more than one stimuli or with an adjustment method. Manipulation of memory load could have resulted in more clear correlation patterns between the perception and the memory conditions.

Since the face is also a complicated stimulus that could hold a lot of relevant social information the perceiving of this information is divided across the face network. If the memory representations are held in the same areas as the perception information, then the memory representations are also distributed across the face network. It could be that since complex stimuli like faces hold more information than simpler stimuli, they are not as efficiently coded from the activity patterns of only one area, due to that the neural signals are

also time-dependent and fMRI signals are slow. For example, the positive correlation of the V1 pattern information to FFA pattern information during the memory delay could be because there is a similar type of information in both areas, but the activity is also time-dependent, so it does not occur as a sustained correlation. This type of time dependence makes it hard to test this hypothesis on fMRI data. Also, the used localizer didn't provide a simple way to separate other face network areas (for example OFA) from other regions of the brain. The localizer also only had neutral faces, so it provided an accurate location for the FFA, but the stimulus used in the perception and memory tasks were emotional faces, and emotional faces tend to activate the face network more widely (Atkinson & Adolphs, 2011).

4.2 Conclusions

In this thesis, the fMRI activity patterns during the perception and memory of faces were compared for the first time. The results indicated that especially in right hemisphere the activity patterns contain face related information during the memory delay. The correlations between the perception and memory RDMs were surprisingly low given that identical stimuli were used in the compared conditions. That could be due to the neural noise that occurs in memory representations or maybe the limited characteristics of the fMRI method. In the future studies, a similar type of analyses could be used to study memory-related processing with bigger sample sizes and different memory tasks

In conclusion, this thesis provided proof that the memory representations for face stimuli are found in the face network. There was a correlation between the activity patterns of perception and memory tasks. Some evidence was found for a different type of activity in the FFA and the V1 during the memory delay. FFA held continuous correlations, but in the V1 the correlations rose with a delay, it could be that the basic features of the face are held in FFA, but the V1 participates in the analyzing of these features. Also, the correlation of the pattern information in the V1 to the pattern information in FFA in the middle of the memory delay could mean that there is some perception information on V1 that is also found in FFA during a memory task.

References

Atkinson, A. P., & Adolphs, R. (2011). The neuropsychology of face perception: beyond simple dissociations and functional selectivity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1726–1738.

Baddeley, A. (1992). Working memory. *Science*, *255*(5044), 556-559.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423.

Baddeley, A. D. & Hitch, G. (1974). Working Memory. In *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic press.

Bays, P. M. (2014). Noise in Neural Populations Accounts for Errors in Working Memory. *Journal of Neuroscience*, *34*, 3632–3645.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science*, *321*, 851–854.

Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*, 641–651.

Christophel, T. B., Iamshchinina, P., Yan, C., Allefeld, C., & Haynes, J. D. (2018). Cortical specialization for attended versus unattended working memory. *Nature Neuroscience*, *21*, 494–496.

Cowan, N. (2001). The magical number 4 in short term memory. A reconsideration of storage capacity. *Behavioral and Brain Sciences*, *24*, 87–186.

D'Esposito, M., & Postle, B. R. (2015). The Cognitive Neuroscience of Working Memory. *Annual Review of Psychology*, *66*, 115–142.

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES-a database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, *42*, 351–362.

Eickhoff, S. B., Paus, T., Caspers, S., Grosbras, M. H., Evans, A. C., Zilles, K., & Amunts, K. (2007). Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *NeuroImage*, *36*, 511–521.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, *4*, 223–233.

Haxby, J. V, Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biol Psychiatry*, *51*, 59–67.

Haxby, J. V., Petit, L., Ungerleider, L. G., & Courtney, S. M. (2000). Distinguishing the functional roles of multiple regions in distributed neural systems for visual working memory. *NeuroImage*, *11*, 380–391.

Jackson, M. C., Wolf, C., Johnston, S. J., Raymond, J. E., & Linden, D. E. J. (2008). Neural correlates of enhanced visual short-term memory for angry faces: An fMRI study. *PLoS ONE*, *3*, e3536.

Jiang, Y. V., Shim, W. M., & Makovski, T. M. (2008). Visual working memory for line orientations and face identities. *Perception and Psychophysics*, *70*, 1581–1591.

Kriegeskorte, N., Mur M. & Bandettini P. (2008). Representational similarity analysis – connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, *2*, 1–28.

Kriegeskorte, N. (2011). Pattern-information analysis: From stimulus decoding to computational-model testing. *NeuroImage*, *56*, 411–421.

Lee, S.-H., & Baker, C. I. (2016). Multi-Voxel Decoding and the Topography of Maintained Information During Visual Working Memory. *Frontiers in Systems Neuroscience*, *10*, 2.

LoPresti, M., & Schon, K. (2008). for Social Cues Recruits Orbitofrontal Cortex and Amygdala: a Functional Magnetic Resonance Imaging Study of Delayed Matching To Sample for Emotional Expressions. *Journal of Neuroscience*, *28*, 3718–3728.

Lorenc, E. S., Pratte, M. S., Angeloni, C. F., & Tong, F. (2014). Expertise for upright faces improves the precision but not the capacity of visual working memory. *Attention, Perception & Psychophysics*, *76*, 1975–1984.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.

Salmela, V. R., & Saarinen, J. (2013). Detection of small orientation changes and the precision of visual working memory. *Vision Research*, *76*, 17–24.

Schobert, A.-K., Corradi-Dell'Acqua, C., Frühholz, S., van der Zwaag, W., & Vuilleumier, P. (2018). Functional organization of face processing in the human superior temporal sulcus: a 7T high-resolution fMRI study. *Social Cognitive and Affective Neuroscience, 13*, 102–113.

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences, 109*, 8780–8785.

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review, 121*, 124–149.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision, 4*, 1120-1135.